

DOCUMENT RESUME

ED 270 843

EA 018 517

AUTHOR Cohen, David K.; Murnane, Richard J.
TITLE The Merits of Merit Pay.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.;
Stanford Univ., Calif. Inst. for Research on
Educational Finance and Governance.
PUB DATE 85
CONTRACT PR-4433
NOTE 46p.; For related documents, see EA 018 515-516.
PUB TYPE Reports - Research/Technical (143) -- Information
Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; Elementary Secondary Education;
Incentives; *Merit Pay; Premium Pay; Success; Teacher
Administrator Relationship; Teacher Effectiveness;
*Teacher Evaluation; Teacher Participation
IDENTIFIERS Evanston School District IL; Midland Public Schools
TX; Virginia Beach Schools VA

ABSTRACT

A study of six selected school districts that had been using merit pay plans successfully for at least 6 years provided insight into administrative strategies associated with merit pay program success. The researchers visited the districts, interviewed teachers and administrators, and studied local documents. Each district used a unique combination of strategies from a list of four: the programs can involve (1) providing extra pay for extra work rather than for higher performance, (2) involving teachers in the establishment of merit criteria, (3) minimizing the impact of awards by keeping them small or distributing them widely, or (4) keeping the program profile low by limiting publicity or making participation voluntary. In the districts studied, teaching quality was generally high, relations between teachers and administrators were good, morale was high, the communities were economically and socially advantaged, teacher salaries were good, and teacher evaluation criteria were broad and not tied to student performance. The merit pay programs examined seemed to encourage work outside the classroom and useful evaluation of teachers, but appeared to have little direct impact on instructional practice. The programs also seemed to encourage community support by adding to the district's aura of accountability. This report discusses three of the districts in some detail and concludes with a discussion of the ironic finding that merit pay programs may be most successful when conditions exist that are likely to lead soon to the elimination of the programs' value--that is, when most teachers exhibit comparably high merit. (PGD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

David K. Cohen and
Richard J. Murnane
Michigan State University, and
Harvard Graduate School
of Education
Winter, 1985

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.

Minor changes have been made to improve
reproduction quality.

Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

THE MERITS OF MERIT PAY*

Merit pay for teachers has been in the news, but it isn't the first time. When Warren Harding was President, many school boards were enthusiastic about rewarding better teachers with more money. By some estimates, between forty and fifty percent of the country's urban school districts adopted some such scheme in the 1920s.¹ This early ardor cooled in the thirties and forties, but revived during the school worries of Dwight Eisenhower's presidency. It cooled again and then revived when President Reagan endorsed the notion a few years ago. The idea seems to provoke more controversy now than it did in the twenties, but many state and local politicians have jumped on the bandwagon.

It isn't surprising to find a business-oriented President pushing merit pay. For advocates of the idea always have argued that it was nothing more than a way to get public schools to follow a sensible business practice: use money to reward the employees who do the best work, and thereby inspire other workers to do a better job or suffer the economic consequences.²

*This is one of a series of reports on merit pay in education. The research was supported by the Institute for Educational Finance and Governance, Stanford University. We were ably assisted by Katherine Jamentz and Niall Nelson. (Contract #PR 4433), which itself is supported by the National Institute of Education. Neither Stanford or NIE are responsible for the views expressed in this essay.

It would be enlightening to trace these enthusiasms in detail, but we are more interested in how merit pay works. What happened when schools tried the idea? Can we learn anything about improving education from their experience?

Such an inquiry might look to some like shooting fish in a barrel. For collective bargaining is now a fact of life in a very large fraction of America's school districts, and teacher unions have been notoriously allergic to pay-for-performance schemes. Even if the idea worked well once, how could it survive today? There also is cautionary evidence from business firms. Though the fans of merit pay hold up business as a good example, many firms do not peg workers' salaries to performance, in part because performance in many jobs is difficult to measure reliably. It might be easy to pay newsboys more if they sell more papers, or linotypists more if they set more type, but what about the people who write the news stories, or design the ads, or edit the copy? As in teaching, it might be easy to distinguish outstanding from terrible work in these departments, but could one reliably distinguish outstanding from merely good work? Figuring out responsible and defensible performance ratings for workers in between the extremes might be difficult at best, and implementing such a scheme could cost an organization much more, in conflict, morale problems, and the like, than the benefits it ever could produce.³

These are worthy worries, but they are no reason to ignore merit pay in education. After all, this idea has been powerfully though intermittently appealing to school boards, administrators, and school

reformers for more than half a century. Despite the cautions summarized here, thousands of districts have tried merit pay. Anything that has such allure is worth looking into, if for no other reason than that it has been taken so seriously by so many people who matter in American education. In addition, the idea makes some intuitive sense. Even if it might be difficult to distinguish between good teaching and better teaching, could it be that hard to see the difference between outstanding and bad teaching? Would not the top teachers be pleased to be paid more than the bad ones? Might it not be good for both sorts of teachers? The good teachers would presumably be better off for the money and recognition, and the bad ones might get the message: improve, or look forward to hard times.

PAST PERFORMANCE

We begin with the historical record on merit pay for teachers -- thin as it is. Lots of school districts adopted some version of merit pay, but between 1920 and 1940, only a few education researchers paid attention to how merit pay worked in the thousands of natural experiments with salary incentives that were underway.⁴ Most researchers studied achievement and ability testing, curriculum, and teachers' classroom performance. These were subjects of a conventional and technical nature, more likely to appeal to education professionals than efforts to use "business methods" in schools. Incentives turned out to be a topic that education researchers disdained for decades -- the entry rarely appeared in journal indexes -- even though it was a concept revered among economists.

But we can learn a bit from practitioners' reports, and the few studies that were done between World Wars I and II. The main impression from these materials is that most merit pay schemes had brief careers, and that some were quite controversial. Why?

The chief complaint was that schools found it difficult to devise defensible criteria of meritorious teaching. The earliest advocates of merit pay conceded that there was little consensus on how quality in teaching might be defined, little evidence for any of the possible definitions, and only weak means for measuring any conception of quality.⁵ But researchers dove into these issues with a vengeance in the 1920s and 1930s, and produced a small mountain of studies designed to identify the components of quality in teaching. This work was a great success in one sense: many different approaches were devised and tried, including everything from supervisors' subjective ratings to detailed research on teachers' classroom behavior. But in another sense these efforts were a dismal failure. Because they measured so many different things, the instruments only reflected the diversity of opinion about what good teaching was, and how it could be measured. The problem of a lack of any measures of teaching quality was supplanted by a Babel of conflicting voices about what the right measure was.

This situation was embarrassing to aspiring education scientists but many school systems simply plowed ahead. Most devised their own ways to decide which teachers were best. Various rating schemes were tried, but few seemed to endure. Some districts relied on principals' ratings, but while this procedure had the virtue of simplicity, many teachers found it objectionable. They worried that "personality conflicts" would be the basis for denying merit increases, and they complained of favoritism.

Other districts used paper-and-pencil rating schemes carried out by central office staff, but there were complaints that these were unreliable and difficult to administer. The historical record is too scanty to know just how much discontent and conflict these problems caused, but the issues were sufficiently important to be mentioned by administrators and researchers who reported on the results of merit pay schemes.

The most problematic feature of all the rating schemes, though, was that none of them could establish a stable connection between teachers' merits and students' performance.⁶ Teachers who had high quality ratings did not consistently produce students' with high achievement. This was problematic, for the great rationale for merit pay was to stimulate productivity. If the better teachers didn't produce better students, why pay better teachers more?

No one knows how influential this problem was, in the demise of merit pay in the 1930s and 1940s. While the research seems to have been fairly well known, we hesitate to give it great weight. Most innovations in education have short careers, rushing in and out of favor for reasons that seem to have little to do with scientific evidence of any sort. The demise of merit pay in the thirties and forties was in this respect typical, not a freakish episode requiring a special accounting. In addition, other explanations seem more efficient. The Great Depression and World War II were good reasons for educators to suddenly lose interest in merit pay. And it seems that many school districts found merit pay a bothersome impediment to efforts to tidy up school administration, and especially to devise simple, unified salary schedules.⁷

One additional reason to be skeptical about the influence of research is that merit pay began to boom again, soon after the war's end. At least

one in every ten public school systems had adopted some version of merit pay by 1960, in a rush of amnesiac enthusiasm. But again, interest quickly dissipated: most of those plans had been abandoned by the early 1970s. In search of explanation, one study reported that the chief local reasons for abandoning merit pay included teachers' discontent with merit ratings and the difficulties of devising a scientifically defensible measure of teacherly merit.⁸

Like some other researchers, we became interested in merit pay in the early 1980s, partly as a result of renewed political attention to the idea. Our first forays into this curious history led us to two guesses. One was that finding any district that had an operating merit pay scheme would be roughly like finding teeth on a hen. We were wrong. It turned out that despite earlier disappointments, there were quite a few districts with working merit pay plans in 1983, when our inquiries began. Some were several decades old. This was a little puzzling, in light of the many previous problems and disappointments, but it was a welcome puzzle. At least we would have living examples to scrutinize, to learn about how merit pay worked, when it did work, and why it had held on in some places.

Our second guess was that in order to last, merit pay plans had to solve three problems that had plagued districts in earlier times. One was to devise a defensible way to decide which teachers were meritorious. A second was to devise ways to minimize conflict among teachers, given the great diversity of opinion about what good teaching was, and the great potential for suspicion and division when some teachers were singled out as being top-flight. A third was to devise a means to cope with the absence of any stable connection between various measures of teachers' quality and various measures of their students' achievement.

PLANS THAT PERSISTED

Our second guess seems to have been better than our first one. We selected six districts that had been using merit pay schemes for at least six years. More than one hundred plans were reported to be in operation, but finding six that seemed worth exploring was not easy.⁹ We asked the superintendents in these districts if they would participate in our study, and when they graciously agreed, we visited each district. Our intentions were ambitious, but pretty straightforward: we wanted to know how merit pay worked, and why, and to what effect. To this end we spoke with scores of teachers (meritorious and not), many principals, and several central office administrators. We studied local documents on each scheme, and pestered local officials with follow-up questions. Our investigations showed that each of the six districts had found a way to cope with the three problems mentioned just above, and it looked as though these coping strategies helped merit pay to endure.

The four most important strategies can be quickly summarized.

- Define the scheme as extra pay for extra work, rather than higher salaries for higher performance at the common classroom tasks of teaching.
- Solve the problem of defensible merit criteria in political rather than scientific terms: make the invention of criteria a collaborative project with teacher representatives; or let each teacher define the criteria for himself; or both.
- Manipulate the merit awards so as to minimize provocation: either keep the amounts negligible, keep the differences among awards small, or pass them out to nearly everyone.
- Keep a low profile on merit pay: make participation voluntary,

and don't make a big deal about who does and does not receive the awards.

Each of the six districts employed some combination of these strategies, though no two used the same combination, and of course none used all of them. In each of the districts merit pay seemed a workable, acceptable part of district operations. In none of the districts did it seem to us, or to our informants, that merit pay was particularly controversial. But one reason for the acceptance and lack of controversy was that the strategies implied redefinitions of merit pay -- often very substantial redefinitions. In historical terms, for instance, the second strategy appears to be a relatively recent invention. Prior to the last few decades, administrators held on to the definition of criteria very tightly; they rarely consulted with teachers, and many tried to solve the problem in scientific terms. Most reports on merit pay between the two world wars also suggest that participation was not voluntary, that most schemes encompassed all teachers. On the one hand, then, it can be said that in our six districts something called merit pay has become a regular fact of school life, accepted and perhaps even constructive. But on the other hand, can such redefined plans fairly be called merit pay? This is a familiar conundrum in the study of organizational innovations, but one that was illustrated with particular elegance in several of the districts that we studied. A few examples will sharpen the point.

The simplest plan that we found was that of the schools in Virginia Beach, Virginia, a self-described suburban district that enrolls about 54,000 students. Educators in Virginia Beach regard their district as a standout in the state. They cite, in addition to good salaries -- average about \$25,000 -- good students, and schools with a reputation for good teachers and serious schoolwork. Merit pay is open to all teachers who

have at least ten years' experience, who have received "entirely satisfactory" ratings on the system's regular teacher evaluation program for the two years preceding their application for merit pay, and who demonstrate "professional involvement." Plenty of teachers meet the first two criteria, 1,366 out of 2,931 teachers in 1982-83, according to the district's figures. But only 540 teachers were earning merit pay in that year, because 826 either did not apply, or were turned down because they did not achieve satisfactory professional involvement.

This low rate of participation seems a little puzzling, given the district's rather relaxed definition of professional involvement. According to the policy guidelines for merit pay (the technical term used there is "Career Service"), Virginia Beach teachers need only complete six activities out of a list of eighteen that includes: successfully completing three hours of graduate study in teaching, successfully completing a school-sponsored inservice education course; serving as a member of a district-wide committee; serving as a member of a curriculum development workshop; accepting a student teacher, and so on. The nineteenth item on the list allows consideration of other professional activities not enumerated, if they were valuable to the school system. Completing six activities on such a list hardly seems arduous. In fact, the list looks very much like a compilation of teachers' common activities, including sponsoring student clubs, and attending conferences related to teaching. This isn't surprising, since the list was drawn up in consultation with teacher representatives. Why, then, do only about forty percent of the eligible teachers participate?

Part of the answer seems to be the size of the award. In return for completing six of the specified activities in two years preceding their

application for merit pay, teachers annually receive the grand sum of \$600. The administrator in charge of the program observed that an award of that size "really doesn't make a big difference."¹⁰ Teachers do get to keep the increment for three years, and then have to reapply every three years until they leave, retire, or decide that the money isn't worth it. Several teachers and administrators told us that the stipend was so small that many teachers simply didn't bother. Another administrator agreed, saying that "if we made the carrot worthwhile . . . [more] teachers would apply."¹¹

The stipend is not as small as it seems, though. For teachers who get merit pay in Virginia Beach move onto the Career Salary Schedule, and thereby gain entry to additional benefits, unavailable to others. A Career Service teacher who earns a Certificate of Advanced Study -- a professional certificate that usually follows an M.A. degree, and requires about the equivalent amount of coursework -- is awarded an additional increment of \$600. That boosts the annual merit increase to \$1,200, on average a five percent raise. A Career Service teacher who earns a doctorate is awarded an additional \$1,200, for a total annual merit increase of \$1,800, -- roughly an eight percent raise. So, for teachers who pursue their own education beyond the master's level, the merit award seems pretty substantial.

But not many teachers pursue their education beyond the master's degree, and the merit awards for such work probably wouldn't even cover the cost of tuition. So even for teachers who want more education, the award may seem too modest to be worth the effort. Still another explanation for low participation is that more apply than are rewarded. Each year several applicants are turned down. The decisions about who

wins and loses are made by a central office personnel administrator -- a somewhat unusual procedure among our districts -- based on the teachers' applications, their evaluations, and on rather perfunctory comments by their principals. But while not all applicants slip through the gate, we were told that only a few were rejected each year, not nearly as many as apply.

A last reason for the relatively low participation rate in such a seemingly easy system may be that the merit awards have little to do with classroom work. Although most of the professional involvement items were related to instruction in one way or another, none were directly related. A good argument could be made, in fact, that most were only remotely related to classroom work. One secondary school teacher commented that "the things that they are looking at in this application . . . aren't classroom related." Asked if her professional involvement improved her teaching, she replied, " . . . the nitty-gritty of the academics I'm teaching . . . no."¹² To the extent that many Virginia Beach teachers disliked anything about the merit pay plan, they disliked this feature. Several derided the scheme as utterly unrelated to excellence in teaching, and others argued that it was inappropriate to call the Career Service awards merit pay. As one former teacher, now a principal, commented, "pretty mediocre people can be career service teachers; you can meet the requirements and not be outstanding in the classroom."¹³ To most teachers, the scheme seemed to be simply a reward for extra work. As one high school teacher who was on merit pay said, when we asked him if the award made him to do a better job, "It depends on what you mean by a better job. Academically? No. But co-curricular activities, extra-curricular activities? Yes. It keeps me doing them, it's the only

thing that keeps me doing them."¹⁴

Few teachers with whom we spoke disagreed with this assessment, though several central office administrators felt the program did reward excellence in teaching. In the teachers' view, and in ours, merit pay in Virginia Beach means extra pay for extra work. It is a rudimentary version of what might be called differentiated staffing. But while teachers were cognizant of this rather special definition, or redefinition of merit pay, few seemed to resent it. Only a handful of teachers said that they thought the scheme was a poor one. Most said that they thought the plan was positive. By way of explanation, several teachers pointed out that participation in the Career Service Schedule was voluntary, and that there seemed to be little competition over the awards. Several pointed out that teachers were little aware of who won. The program had a very low profile, and release of the winners' names actually was forbidden by district policy. We also were told that the scheme, which dates to 1974, was created with teachers' participation, and has been modified somewhat during the intervening decade in response to their comments. All of these conditions seem to have reduced the chances of controversy, and thus increased teachers' acceptance of the scheme.

In fact, the most striking feature of teachers' comments on merit pay was their very positive assessment of the district and its relations with teachers. When we asked why they liked working in a district with merit pay, they all gave roughly the same answer -- Virginia Beach was a classy system. One teacher pointed out that administrators "back teachers emotionally . . . they let you know that you are doing a good job. This system is a good system . . . [and] staff morale . . ." is high. Merit pay was just one example of community support for good education: "There

is a great push to succeed here. The kids demand a good education."¹⁵ Another teacher expanded on the point: "There is a tradition of good education . . . this is one place where teachers are dismissed if they are not good . . . in this district you have to support quality instruction."¹⁶

For these teachers and many others, merit pay was acceptable and even positive because they saw it as one manifestation of the district's support for good teaching. The Career Service was an admittedly modest manifestation, less important than several others, but it was consistent with the district's general orientation to quality. That made Virginia Beach seem uniquely desirable among the other local districts that teachers knew. So while it seems fair to say that merit pay works in Virginia Beach, it also seems that merit pay works there because it modestly compliments other educationally positive features of the school system.

Another way to put the point is that merit pay made sense in Virginia Beach precisely because it blended easily into the woodwork. It was a weak treatment, successful partly because of that fact. The plan's modesty contrasts sharply with most of our other districts. Perhaps the most complex and politically prominent scheme was employed in Evanston, Illinois, an affluent suburb of Chicago. This district has had several versions of merit pay since the mid-1950s, and the changes bear directly on our story.

Initially, we were told, merit pay ". . . was pretty much a question of what the principal thinks . . . and if he liked what you did you got merit pay. There was no formality."¹⁷ That system lasted about a decade, and then was radically overhauled. The result, installed in the

mid-sixties, tied a newly devised set of criteria for good teaching to a rather elaborate arrangement for teacher salaries and promotion. There were five salary tracks, ranging from an initial track for novices to a fifth (and top) track for the most accomplished professionals. Each track had a dozen or more salary steps keyed to years of service, and teachers in each track automatically moved up a step every year. Doing so earned them annual increases that ranged between about \$200 and \$800, averaging about \$500. But teachers could not move from a lower to a higher track simply on the basis of seniority, and salaries were higher in the higher tracks. Moving from one track to the next highest track (at the same step for years of service), would produce salary increases that ranged roughly between \$1,000 and \$3,000. And to get "track movement," as it is called in Evanston, teachers had to achieve a designation of "meritorious" under the new criteria.

Deciding what is meritorious is of course a long-standing sore point in merit pay, but Evanston's solution seemed ingenious. Administrators, teachers, and an outside consultant worked up a list of five dimensions of meritorious teaching. Teaching skills were at the top of the list, followed by "classroom and school environment," "communication," "interpersonal relationships," and "professional contributions." A list of "competencies" was then set down within each of the five areas. Teaching skills, for instance, began with "exhibits and imparts knowledge of subject matter," moved on to "encourages each member of the class to participate," and then to "uses the curriculum guideline in a creative and critical way." These and others seem sensible points, and several would not even be mentioned in other districts' evaluation schemes -- if they had such a thing. Evanston is a very well-regarded district, with a

long-standing reputation for quality teaching. It shows through the professional jargon.

The next category (classroom and school environment) seemed a little less specific, but drew attention to classroom discipline, constructive use of time, and related issues. Communication and interpersonal relations focused on such vague but noteworthy matters as working cooperatively with colleagues, and "two-way communication" with parents. Professional contributions was a list of mostly familiar items, such as supervising student teachers, participating in curriculum development, taking "professional leadership," and the like.

The general drift of these criteria was plain enough: classy teaching, community service, good relations with parents, and professional work beyond the classroom. Nonetheless, there still was plenty of room for differences of opinion about what each criterion meant, and whether a given teacher had met the mark. Evanston's 1965 plan tried to solve that problem by avoiding a search for scientific truth in favor of a decision-making process based on professional knowledge. Teachers who wanted track movement had to apply for it, and each applicant had to document his accomplishments. Applicants were enjoined to "be specific with details, dates, and all necessary information that will clarify . . . what you have done."¹⁸ When teachers had documented their meritorious work, their applications would go to district advisory committees composed chiefly of teachers. The committees (one for elementary and one for middle schools) would read the applications and make the decisions.

Here was a process that legitimated decisions about merit pay by delegation to professionals. Initiative and presentation of evidence rested with individual teachers, and decision-making rested with a

committee that was selected by lot from a pool of possible teacher representatives, all of whom had been elected by teachers. Teachers had to figure out what meritorious teaching was and who should get it, subject to guidelines that had been devised by teachers and administrators. It was a serious procedure, and clever problem solving.

But if this arrangement solved some problems, it gradually seemed to create others. One problem arose from the vagueness of the concept "meritorious," and the injunction to "be specific:" for the applications quickly escalated in size and detail. The teachers' documentation became "dissertations . . . the system became cumbersome."¹⁹ "The application . . . could turn out to be . . . fifty pages long."²⁰ Such a monument to oneself was not easy for many teachers to build, for it took time, required careful record keeping on one's work, and also seemed to require that teachers brag. A second problem was that some applicants were better at producing a convincing document than at teaching a good class. One long-time staff member said that ". . . if you could write a good application you got it," quite apart from whether you were a good teacher.²¹ Another problem, less predictable but delicious, was that the teachers' committees that passed on applications turned out to be quite conservative. Instead of rolling logs and trading pork, "the teachers were very hard on each other." As a result, it became difficult to get track movement.²²

So, an ingenious process for managing a difficult set of issues produced a new set of problems. In a district with lots of good teachers, lots of good teachers did not get merit pay. As time wore on, discontent mounted. That created pressure for yet another revision. The school district administration gave an additional push by initiating a new system

of teacher evaluation, in the middle-to-late 1970s. The administration's reasons for the new system appeared to have little to do with changing merit pay: it simply wanted a better way to assess all teachers' performance. Administrators worked in close consultation with teachers and teacher union representatives, and produced a procedure that had two main parts. One was annual evaluation in terms of specific job targets jointly selected by each teacher and his or her principal; the other was a general performance appraisal which ranged over much broader and more diverse issues. The only requirement in the first part was that targets always include some that bore on student performance -- though students' actual achievement was never considered. The general performance appraisal, however, drew heavily on the mid-1960s merit pay plan. It set out a list of evaluative criteria that were nearly entirely taken from the old criteria of meritorious teaching. And it specified that teachers should write a narrative -- seemingly a short form of the old merit application -- reporting on their performance within those categories. The job target evaluation, by contrast, was carried out by principals, and it always included observation of classroom teachers.

The evaluation scheme seemed to balance administrative and teacher interests reasonably well, and it contained a new scoring arrangement. Teachers would be awarded one of four ratings: exceeds expectations; meets expectations; needs to improve; and unacceptable. After the new evaluation plan had been in operation for a few years, Evanston teachers decided that it should be substituted for the old merit pay application procedures. While teachers have some problems with the new evaluation -- chiefly large variations among schools in the criteria on which principals chose to focus, and in the severity with which they apply their favorite

criteria -- they had more problems with the old merit pay system. As a result, in 1982 a new merit pay arrangement came into operation, which tied track movement for teachers to their ratings on the new evaluation.

At first glance this looks like a lovely marriage of efficiency and participatory government. But matters were not quite that simple. One qualification is that the old merit pay criteria, carried forward into the new general performance appraisal, embodied a diverse list of criteria for merit. Classroom performance topped the list to be sure, but professional accomplishment and community service were amply represented. All are desirable qualities in a school system, and perhaps even essential for good schools, but one effect of including several criteria of merit was to dilute the importance of classroom work. For it added extra pay for extra work onto the idea of extra pay for excellent teaching. This is not a bad idea from many angles, but it is not the first thing that comes to mind when merit pay is mentioned.

Another qualification is that excellent teaching is not the only way to get track movement in Evanston. Teachers who get scores of "exceeds" can move across tracks more quickly, but teachers who simply "meet expectations" can still move across tracks: it just takes longer. In fact, the only teachers who cannot get track movement are those rated "unacceptable," or "needs to improve." As one teacher with many years of service in the district told us: "The current system was initiated by teachers because they wanted the movement across tracks to be automatic . . . this would make it more accessible."²³ He went on to point out that anyone can get to the third track simply by getting "meets expectations" ratings. Moving beyond that to the two top tracks could be done by combining those same scores with a masters degree, some graduate

course credits beyond the MA, and a good deal of community and professional service.

This certainly means that teachers would have to work hard to get to the top, and that they would have to be good teachers. It is also plain that teachers who got the top ratings would rise faster. But being tops in the classroom is not required to get to the top of the scale. One could do it by simply "meeting expectations," by doing a good deal of community and professional service, and by persisting. Teachers in Evanston, by the way, consistently told us that in their view, "meets expectations" was a grade of "C."

Not surprisingly, Evanston teachers have a generally positive view of these arrangements. For one thing, they have an appreciable piece of the action. They had a large hand in devising and revising all parts of the plan, and they made out pretty well financially. At the end of the 1983 school year, for instance, something in excess of 55% of all the teachers were on track three or above, and nearly a third were in the top two tracks. Or, to slice the cake a little differently, the average salary in the district was about \$26,000, and roughly a third of the teachers were earning more than \$30,000 annually. Teachers did not regard these results as top-heavy, given that Evanston has long recruited unusually good teachers. And they reminded us that the salaries were not great, considering these teachers' experience, their job requirements, their quality, and preparation. But they are good salaries as teaching goes. Teachers and administrators did tell us again and again that the Board of Education worried that too many top ratings had been handed out. It also may trouble some officials that, as one teacher put it, track movement for competent teachers is pretty much immune to administrators' views. Even

if principals don't think a teacher is great, "you can always progress, even without high administrator evaluations, if you work at it long enough."²⁴

But if we step back and ask if this story satisfactorily explains why Evanston teachers had a mostly positive view of merit pay, we would answer in the negative. We would have to add that teachers there seemed to like merit pay chiefly because they saw it as part of the district's emphasis on quality education. One teacher told us that "parents have a profound effect on teacher performance here. They expect good teachers, excellent teachers, and they aren't hesitant to pass that expectation on to the teachers and to the Superintendent."²⁵ Like everyone else, she knew that the district had hired and kept good teachers. Merit pay was only one modest reflection of that reputation. One administrator, speaking of the district's teachers said ". . . we work hard to screen people out. I have to thank the teachers' association for much of that. They don't fight relentlessly for a bad teacher. They check to see that we are doing everything legally, of course. But they want to see bad teachers out too." He also suggested that the values embodied in the merit pay plan did not arise from the plan, but from long-standing community traditions. On teaching, for instance, he attributed quality to ". . . the composition of the community . . . there is almost a haughtiness to it When you have that kind of community, eventually you get a different kind of teacher." And in discussing community involvement and professional service, he noted that ". . . it's almost impossible to live here without being involved. You can't live here without someone coming to your door and asking you to work" The criteria for track movement simply reflected community values and practices: ". . . because of the kind of

community this is, teachers don't see it [service and involvement] as unusual."²⁶

These comments closely resembled what we heard in Virginia Beach. There were many important differences in the two districts' merit pay plans, most notably the prominent role of the Evanston scheme compared with the very modest importance of the Virginia Beach plan. Despite these differences, when teachers tried to explain why merit pay made sense, their comments ran in parallel tracks: they saw the schemes as part of a broader commitment to quality in education. Merit pay struck them as an expression of that commitment, and perhaps as a modest reinforcement for it, but not as a basic cause of quality. So, Evanston teachers, like their colleagues in Virginia Beach, support merit pay for what many fans would regard as the wrong reason: because merit rewards reflected community esteem for good teaching, not because they seemed a strong or effective way to improve teaching.

In retrospect it does not seem surprising that community traditions should be at least as important as design features, in explaining why merit pay appears to work in these communities. But then, in retrospect, nearly anything in social science can seem unsurprising. Before gaining the advantage of hindsight we had not expected local traditions to loom so large, nor has this point been much noticed in either the research or the prophetic literature on merit pay. We stress the point because it throws a somewhat novel light on how merit pay works, and why.

In this connection, consider another of our districts -- Midland, Texas, a system of about 17,500 students that is situated in a small and rather well-to-do oil and financial center. From a design perspective, Midland would seem to have committed a major faux pas, since less than six

percent of its eligible teachers have won merit pay, even though the plan provides rather sizable rewards and has been working for more than half a dozen years. Midland has a four-track teacher advancement plan, much like Evanston, and advancement across the tracks is keyed to a very similar evaluation scheme. Good classroom evaluations are required for advancement to any track, but making it onto the top two tracks also requires considerable professional and community service. Teachers who make it onto the first track, which means that they are no longer mere probationary novices, receive only a \$1,000 annual increment. But moving to track two produces \$2,000, to track three \$3,000, and to track four \$4,000 annually. The awards stick for four years, and then re-application is required.

This seems like a tidy little package of political dynamite: ample rewards, closely tied to teachers' status within the district, and all bound together in a rather elaborate package of teacher evaluation. Yet there are very low participation rates. Teachers in Midland nevertheless gave the plan favorable reviews. How can that be explained?

Part of the answer does seem to lie in the plan's design. First, the evaluation criteria in Midland cover the same three bases as those in Evanston: classroom performance (observed by principals), professional accomplishment, and community relations. Merit is not exactly detached from classroom work, but the importance of classroom work is reduced by the inclusion of other values. For instance, teachers cannot achieve level II without earning an M.A. degree. They cannot achieve levels III or IV without doing a good deal of community and professional work outside the classroom, and earning either fifteen (III) or thirty (IV) credit hours beyond the M.A. Teachers who aspire only to excellence in the

classroom are thus effectively barred from the top three levels of the Midland system. One way of looking at this requirement is as a way to broaden merit criteria so as to sweeten the pot for teachers, and take the heat off classroom performance. But another way to look at the education requirements is that they permit good teachers -- or any teachers, for that matter -- to remain outside the merit advancement system without raising any questions about the quality of their classroom work. For they can point to the educational requirements as the reason for their not gaining the salary increments. As in all the districts that we studied, student performance is not mentioned.

A second design element that helps to explain why merit pay works in Midland is that teachers seem to have played an important part in devising the scheme. They served with administrators on a mid-1970s committee that reviewed reports on merit pay in other districts, and they helped to work out the evaluation criteria and procedures. They also serve with administrators on the committee that reviews and decides on teachers' applications for merit pay. Merit pay works in Midland, partly because the scheme has been politically legitimated by teachers' participation, as in Evanston. In addition, the Midland program seems to have quite a low profile -- unlike Evanston. Teachers seem uninterested in who is on or off the merit ladder, and some seemed unaware of details of the plan. While the district announces the results, it does so in a distinctly low-key manner, in the board of education minutes.

Finally, to win merit pay, Midland teachers must apply and must document their claims. And despite local administrators' insistence that much of the decision about merit rests on the principal's observations, documentation seems to loom large in both the application and in the

results. The district's guidelines remind teachers that their application "should be based largely on documenting your own individual [sic] performance as a teacher." It requires "signed statements from your departmental chairperson, principal, coordinators and/or support personnel who are in a position to observe your work . . ." to support such claims as "shows evidence of effective lesson planning." And it enjoins applicants to document their performance on each criterion "as specifically as possible."²⁷ Local informants disagreed about the impact of these requirements. Administrators responded negatively to our questions about the possibility that the documentation requirements would reward skill in presentation and organization, rather than in teaching. But several teachers took the opposite view. One told us that, in her view, presentation was very important, and that "being able to type" gave her a real advantage.²⁸ Another teacher, who was a strong ideological partisan of merit pay, lost out in his first several applications. Upon inquiring about the reasons, he was led to suspect that he "had turned in a pretty skimpy amount of documentation." He had assumed, he said, that "they know me and what I'm doing." But after the rejections he took another tack: "I pulled out all the stops. I really played their game and went overboard, I guess. I documented things that I thought were asinine I think I finally turned in a notebook with something like 257 pages."²⁹ He got merit pay.

Each of these design features helps to explain why merit pay works in Midland, despite the fact that only a few teachers get the rather good awards. But there are some other, rather different explanations, as well. First, when we asked why so few teachers applied for merit pay, and why so few seemed to worry about not getting the awards, administrators

and teachers reminded us that Midland is an affluent little city. One top official pointed out that the financial incentives for applying were not that strong, considering the economic context: the district's salary scale is one of the highest in the state, and many of the teachers have spouses who work in oil or finance. More money was said to be much less important to them than good working conditions. In fact, he pointed out, a disproportionate fraction of those who were on merit pay were divorced mothers. Second, Midland schools had a very solid reputation for quality, and teachers knew it. They liked working in a district with good students, a tradition of support for education, and a good name.

Teachers thus had at least two big reasons to be pleased that they were working in Midland, quite apart from merit pay. From this angle, the very low participation in Midland looks much less provocative than we first thought, partly because merit pay looks much less prominent than we first thought. General affluence surrounding the schools reduced the salience of merit increases within them, and the community's strong support for education provided plenty of non-monetary reasons for teachers to like working there. In Midland, as in Evanston and Virginia Beach, one reason that merit pay seems to work is that it blends easily into the local woodwork.

A cynic might say that this is just a fancy way of admitting that our districts are places in which nearly everything seems to work well. If all we have found is that merit pay works when the living is easy, who cares? While it is useful to know that this scheme can work in places that have all the advantages, doesn't that tell us something about what would happen if merit pay were tried in places that had fewer advantages? After all, we did not find any current examples of merit pay schemes that worked in

districts that enrolled mostly disadvantaged students, or in districts that had poor quality schools, or in districts that had fractious labor-management relations. Nor did we find evidence that merit pay worked in large school districts.

There is something to this view: if we take past practice as a guide, merit pay seems unlikely to work in many American school districts, among them the districts most in need of improvement. But one purpose of our study was to understand what implications successful merit pay plans might have for school improvement, so this limitation seems useful, and timely to report. Even apart from that point, though, the cynic's objection is a little too simple. For past practice is never a perfect guide. It is possible that some less advantaged districts could capitalize on the design experiences of the schools that we studied, and successfully implement merit pay plans. For another thing, professionals in our districts put a good deal of time and thought into devising artful versions of this old idea. Had they been a bunch of bumbler, merit pay almost certainly would have blown up in their faces -- even though community support for teaching was strong. Such pro-education districts, after all, are notable for having more, and more critical citizen participation in school affairs than most local education authorities.

So, we sum up our argument thus far in terms of a conjunction between good design and community values. Merit reward plans for teachers seem capable of surviving

--when communities are largely advantaged, economically and socially, have decent salary schedules and good labor-management relations; and

-- when the criteria of merit include a broad conception of

teacher performance, when they avoid student performance, and when the scheme operates within a carefully crafted framework in which the definition of merit and the process of deciding who wins and loses are legitimated by political and professional rather than scientific means.

We do not advance these as prescriptions for success, but as generalizations from the experience of a rather special set of school systems. Other districts might succeed if they only met some of these criteria.

THE MEANING OF SUCCESS

While we are issuing disclaimers, we might mention one more. The analysis above is neither an endorsement of merit pay for teachers nor an exhaustive account of its impact. Readers already will have noticed that we have so far employed a rather straightened criterion of success, considering only whether the schemes were politically viable. We have explored whether plans were implemented: Did they last, or did they vanish in clouds of angry accusations? If they lasted, were they well-regarded, or did they become a bone of bitter contention? These are not trivial points, as those who have either observed or participated in efforts to improve public education will know, but they are hardly the whole story. For innovations can succeed in any or all of these senses without seeming to succeed in any other sense -- such as actually improving education. Not a few academic kibbitzers on the late Great Society amused themselves for quite a few years, telling just such success stories. Is that the sort of story we are telling here? Did merit pay improve teaching in our districts?

The answers are a little complicated. First, no one told us that merit pay itself improved teachers' classroom work. Several teachers said that the plans discussed here affected instruction, but it was the evaluation, as much or more than the raises, to which they pointed. And even these teachers saw only marginal changes. Nearly all the teachers regarded merit pay simply as recognition for good teaching. One teacher in Evanston echoed the views of many others when she said, in response to one of our queries about the effects of merit pay, "it's nice, it lets you know there is something out there for your work." 29a

It is of course possible that most teachers missed the point, that money rewards for good teaching encouraged better teaching but did so unwittingly. Humans' testimony on their states of mind, and on the effects these states have on behavior, often is unreliable. We cannot rule out this alternative, but we are inclined to be skeptical. One reason for our caution is that so many different teachers made the same point, all quite independently. Another reason is that many teachers freely admitted that merit pay had an effect in other areas. Quite a few testified, for example, that the extra pay for extra work was what kept them involved in professional activities and community service. We could think of nothing that would explain teachers' recognition of an incentive effect in one area and their lack of recognition in another. Nor could we think of a reason to accept one set of teacher testimony while impeaching another. Someone might argue that because teachers' classroom work is so central to their professional identities, they would be unwilling to acknowledge influences on that aspect of their performance. But this view was directly contradicted by several teachers who praised the evaluation systems in several of the districts. They said that the observation and

analysis of their classroom work helped to improve it.³⁰ We can think of no reason why teachers would acknowledge one sort of influence on classroom work and not another.

Another reason we are inclined to doubt that fiscal incentives had an appreciable or consistent effect on performance is that teachers had quite a plausible alternative account of the incentives for performance in their trade. When we asked teachers what motivated them to do good work, their answers were all consistent with what one woman in Virginia Beach told us: "Every once in a while the light bulb goes on in a kid's head."³¹ Another likened her work to the theatre: "Children are a great audience, you can see if you are boring them to death . . . when I hit, it feels good, its like acting, you are on stage, they either make you or break you."³² Another teacher summed up these ideas nicely when he told us, "student support is the biggest motivation. The harder they try the harder I try."³³

Each of these teachers was calling attention to the interactive nature of their work, and to the way in which incentives for performance were therefore anchored in social interaction. Teachers' work consists largely in efforts to improve other humans, or to help them improve. Teachers cannot succeed unless the people on whom they practice try to do well in pursuit of the teachers' objectives. Unlike surgeons or dermatologists, teachers need active cooperation and engagement from their students, not just passive acceptance of a treatment. Students therefore hold the keys to teachers' success. When students try hard and do well, they make it possible for teachers to succeed. If students don't make the effort there is no way that teachers can score an instructional success. Students therefore control some of the most powerful incentives for teachers'

performance. This is part of what many teachers meant when they told us that the motivation for their work was "intrinsic," rather than being monetary. Money is important to teachers, but money is not the key payoff for success in the classroom. The largest rewards are achieved through others' performances.

This perspective on teaching helps to explain why so many teachers prefer students who work hard, especially those who also do well. In addition, it helps to explain why, when teachers are asked about ways to improve their work with students, they don't suggest money incentives. Teachers would like higher salaries, but few would say that their teaching would improve if they were paid more. They do say that better salaries might help to recruit and hold more good teachers. But when asked how to improve their work, teachers point instead to the things that they regard as direct helps in their instructional interactions: fewer students, better materials, more time to prepare and to correct lessons, and the like. Teachers can see how such changes would help them and their students to "turn the light on." Of course, none of this is mutually exclusive with the idea that money incentives would help, but it tends to support the view that there is at least a hierarchy of incentives for teachers' skilled performances, and that money is not near the top of the list.

One last reason we are inclined to accept teachers' testimony on these points is that it was generally supported by management. Few administrators claimed that merit pay improved teachers' work in classrooms, even though they had the strongest reason to make such claims. The superintendent in Midland, for example, was very positive about the system he had helped to install in that district. He thought

that the incentives for professional activity and community involvement had worked. He thought that teachers had benefited from the "self-examination" associated with the evaluation and documentation requirements. He even thought that those teachers who had been promoted "refined their performance" as a result of the self-examination. But he emphasized that the merit increases were rewards, not incentives. And he said, "these were good teachers already -- this program didn't make them good teachers."³⁴ Additional, he made no claim that the rewards had inspired better work from other teachers.

Our conclusion on the educational impact of merit pay is therefore mixed. On the one hand, there is no evidence that the money had an appreciable or consistent positive effect on teachers' classroom work. But on the other hand there is a fair amount of evidence that it did encourage some teachers to do the sorts of work outside the classroom that many educators consider important. There also is some evidence that the evaluation requirements associated with some merit pay plans were useful to some teachers -- though evaluation easily can be, and most often is decoupled from merit pay. In addition, there is quite consistent evidence that these merit pay schemes helped many teachers to feel a bit better about their work, because the schemes seemed to express support for good teaching.

All of these are relatively positive effects of merit pay. Even if none of the effects are central to education, and even though no-one claimed that they were large effects, they run so cleanly across the grain of the worries mentioned by opponents of merit pay -- conflict, favoritism, demoralized teachers, and the like -- that they seem particularly worth mentioning. Of course, it is also worth recalling our

cautions about the rather special nature of these districts.

We noticed one other effect of merit pay. The schemes helped school administrators to win school board support for spending money on education, and it helped both boards and administrators to make the same argument to their communities. One administrator in Midland told us that he didn't think that the teacher incentives were the most important feature of that district's merit pay plan. "I would describe it as an attempt on the part of the school board and administration to have some type of device to justify paying teachers extra money." When asked why such justifications were needed, he said: "I think the demand is coming from school board members and patrons of the schools. Especially when teachers go in and . . . ask for additional pay, they want to know how the money is being used."³⁵

On this view, then, merit pay was important in political transactions over education -- especially school finance and teacher salaries. This phenomenon was not confined to the ideologically more free-market oriented states of the Sunbelt. One of the other districts that we visited was Niskayuna Central, a small, white-collar, suburban district in upstate New York's Mohawk River valley. Niskayuna has had merit pay for roughly thirty years, nearly since the district was first organized, in fact. Everyone there seemed to be happy with the plan. There is a teachers' union in Niskayuna, and the schools have a strong reputation for quality. Community support for education seems high. Despite the seeming smoothness of life in this district, professionals thought that merit pay was important in the local politics of education. One high-level administrator told us that the scheme ". . . does give the school board and the community the feeling that they've got some control over the

staff. There is an evaluation process here and the process can be used in a way which will reward the good and punish the wicked . . ." 36 In Niskayuna and other districts, merit pay helped board members and voters to believe that the district was taking a responsible and even tough managerial line, and that under such circumstances spending additional money on education was defensible.

These points lead to another interesting conclusion about the educational effects of merit pay. At least in some of our districts, the plans seem to have been most useful for their political symbolism. For merit pay is one of those rare political devices that can broadcast quite contrary messages at the same time. It seemed to tell teachers that their work was valued while telling elected officials and constituents that management was taking a tough line with the hired help. It isn't every school improvement scheme that can serve such diverse purposes; it is thus little wonder that school people in these communities seemed so attached to merit pay -- even if no student's achievement improved.

Under some circumstances, then, limited but hardly trivial, merit pay does appear to some have some educational merits. But these seem to lie more in the domain of political symbolism than classroom performance. This is an important realm, to be sure, and one in which many issues that do bear on education, even on classroom work, are decided. It just happens not to be the realm in which the fans of merit pay have advertised the merits of their pet scheme. The districts that persevered with the idea turned it in quite a different direction, in this as in other respects.

DECAY OF MERIT SYSTEMS?

Thus far, the story we have told is a generally sunny one, even if it

has some unexpected twists. One theme in our analysis has been that when merit pay for teachers works, it does so because of a happy conjunction of human craft and social circumstance. Our communities had thoughtfully designed plans for rewarding teacherly merit, but they also had values, finances, and school staffs which made it relatively easy for the merit pay plans to work constructively. But if this happy interaction helps to explain the success of merit pay in some districts, it also seems to be responsible for a curious deterioration in some plans. For several of the districts that we visited have grown progressively less able to maintain distinctions between meritorious and non-meritorious teachers, in part because so many of the teachers were so good. This is an ironical point, but an important one: the relatively high quality of education that made it possible for merit pay to work apparently also tends to defeat merit distinctions among teachers, over time.

What are the dynamics of this curious situation? Part of the answer is that in good schools merit pay puts administrators in a bind. For the good teachers in such schools don't see much distance between themselves and the very good teachers. They therefore get upset if the very good teachers are assigned a higher status, and given more money to boot. From the school managers' perspective, the generally high quality of their staff -- which makes them a pleasure to work with -- makes merit pay a potential headache. We were given rather vivid testimony on this point in Niskayuna. While teachers there seemed quite happy with merit pay, school administrators told us that they had encountered difficulties in making distinctions between outstanding teachers and those who were merely quite good. For instance, one building administrator told a story of rating a teacher as "outstanding" (the next to top rating), rather than

"exceptional" (the top rating). The teacher walked away and said, "If that's all you care, that's all you get." According to this principal, "she defined a new role for herself, did not attempt to provide additional evidence of work She has determined that there is more to life than teaching, and has gone right on her merry way Her classroom performance wavered after the evaluation . . . now it is as good as it was, but it has not improved. . . . I don't think her potential has been reached. I point directly to the merit."³⁷

In this principal's view, making distinctions between very good and merely good teachers was counterproductive, for it created resentment and alienation rather than firing up the good teachers to do even better. Given the skewed distribution of teacher quality in his district -- a high average, and a rather tight dispersion around the average -- merit pay would be a problem if not managed very carefully. Principals in other districts took a similar view. None of them reported any problems in giving low ratings to poor teachers, or to teachers in need of considerable improvement. But many principals reported difficulty in making money and status distinctions among large majority of their teachers, who were all at least good. One building leader in Evanston reported his way of managing this problem -- a device that seemed common in these two districts. He said that "I like operating on the positive. I think giving people the "exceeds" [the top rating] . . . makes them feel good, and when they feel good they do more."³⁸ His approach to managing merit pay, then, was to avoid absolute accuracy within the large pack of good teachers. Instead, he slightly inflated the ratings of the good-but-not-tops teachers. The inflated ratings were at least a defense against discouragement and jealousy, and perhaps, as he claimed, also a

boost to teachers' performance. If these principals were correct, there is a nice twist here: in some cases, honest merit ratings would create disincentives for teachers' improvement, and dishonest ratings would be more likely to produce positive effects.

If many building administrators followed this practice of grade inflation, then over time, ratings would tend to cluster at the top. This is what seemed to happen in Evanston and Niskayuna. But it is worth noting that a skewed quality distribution is not the only source of this result. For one thing, the principals' managerial position is more vulnerable than that of many other executives: principals depend on teachers more than teachers depend on them. Teachers, after all, work in their own self-contained classrooms, and they often can get along decently even if their principal is a boob. A good principal can help teachers, of course, but good principals are not required for teachers to do their jobs. Principals, by contrast, need teachers who do good work if their school is to run well -- and if the principal is to be seen as doing a good job. A pack of poor teachers can probably do more to wreck a principal's working life, and perhaps his reputation, than poor principals can do to damage teachers' work and reputations. Because good principals know this, they try hard to avoid upsetting their good teachers. Needless to say, giving some good teachers merit awards while withholding them from other good teachers, could ruffle a few feathers.

An additional complication is that most teachers in most school buildings are now tenured, and have been for the duration of the merit pay plans that we studied. As a result, barring incompetents, principals have not had many alternatives to the staff members already there. This constraint puts a good deal of pressure on principals to avoid alienating

teachers, since the teachers are unlikely to go away, even if unhappy. For another thing, there is not a long line of talented applicants at most schoolhouse doors. Recent developments -- a surplus of teachers due to layoffs and enrollment decline, shrinking teacher income, and criticism of schools -- have reduced applications. Good applicants seem few and far between. One elementary teacher said that her student teachers in recent years had been of "low quality," and getting worse.³⁹ A high school principal told us that of all the student teachers who had worked in his building in the preceding two years, only one was worth hiring. He noted that many application letters are so full of misspellings that he simply throws them away. In his view, and that of many colleagues, "We're not getting the horses in teaching."⁴⁰ Even if tenure had vanished yesterday, then, the absence of many talented replacements would push principals to husband their existing staffs. But the combination of tenure and a shortage of talented replacements powerfully presses school managers toward school improvement strategies that will not offend teachers who already are doing reasonably well. In Niskayuna and Evanston, principals re-defined merit pay so that it became such a strategy.

This part of our story is full of lively implications. First, there are a few points about incentives. One is that merit rewards seem to be sensitive to the existing distribution of merit. This is not a point that we have found anywhere in the literature on merit pay, and it strikes us as important. If high average quality, and a modest dispersion of quality around the mean tend to defeat individual incentives for quality performance, would the same thing be true in populations with low average quality, and a similarly tight dispersion? Would merit pay work better in

populations with a larger dispersion of quality, and a more average average? Or are inflated ratings and consequent decay of merit distinctions an inevitable response, in any distribution of quality, of managers who want to avoid trouble?

Our districts offer no unambiguous answers to most of these queries. On the matter of decay, Midland has a very low participation rate, and that districts' restrictions on access and general affluence make it likely that participation will remain low. But if there seems to be no danger that merit pay will expand toward universal coverage in Midland, the price of this solution seems to be a program which is quite marginal to district operations, and in which only a few teachers get beyond the first step. If one were to draw a moral about the decay of merit pay plans from the Midland story, then, perhaps it would be that in high quality districts, program insignificance is the price of avoiding a drift toward universal coverage.

Another of the districts that we visited -- Catalina Foothills, an affluent suburb of Tucson, Arizona -- sheds a complementary light on the matter. This is a small district, with about 85 teachers in three schools that enroll students whose parents are mostly professionals. Test scores are said to be second highest of all districts in the state. Five or six years ago a new superintendent was hired, apparently in part because test performance had wavered. He fired two of the school principals, hired a new Assistant Superintendent, and set out to clean house. The Assistant Superintendent said, ". . . we have released 21 teachers over the last five years. We are literally weeding out the bottom ranks of our teachers. And at the same time we are recruiting really good teachers."⁴¹ Soon after he arrived, the new superintendent also

instituted a merit pay plan, aimed at rewarding good performance among teachers. The plan is rather limited: the awards are generally modest (reported to be between \$300 and \$1200 annually), and teachers are strongly encouraged to use them either for professional development (attending workshops, etc.), or for classroom improvement (additional materials or equipment). The awards are made annually, and must be renewed from year to year.

The result has been what the Assistant Superintendent termed "an interesting problem." Due to the recruitment of bright young teachers -- Catalina Foothills is the only district in the state that recruits out of state -- and firings of poorer teachers, ". . . we see the range of performance [among teachers] compacting, and actually moving up to the higher end of the continuum. And the bottom is definitely moving up."⁴² As a result, the proportion of teachers receiving merit pay has grown: at the end of four or five years, nearly every teacher now working in the district seems to have received a merit award. One top administrator estimated that the percentage of teachers who have not participated was "very small."⁴³ The other concurred, saying that ". . . most of the teachers" were in the program.⁴⁴ So, in Catalina Foothills, improving the quality of the teaching staff seems to have helped push merit pay toward universal coverage. This story seems to be consistent with the drift of our analysis: the more uniformly good a teaching staff is, the more difficult it will be to maintain quality distinctions, and the more likely that merit pay will drift toward universal coverage.

Our story also suggests another point concerning limits on merit incentives: they are sensitive to the circumstances in which they

operate, as well as to the distribution of quality. Labor markets are one such circumstance: when there is no queue of qualified replacements, performance incentives will be weakened. When most employees are tenured, money incentives for performance may be ineffective, or even counterproductive. Many commentators on education seem oblivious to these points. They often view merit pay as a counterweight to the evils of tenure, but our investigation makes it plain that merit pay easily can be problematic under conditions of tenure, degrading rather than improving teachers' work. A last point about constraints on merit incentives is that constraints can interact. We found a relatively high and tight distribution of quality, combined with tenure and signs of weak replacement quality. The combination tended to make school managers cautious about merit pay. Their caution strikes us as rational.

These observations may have application beyond education, or merit pay, or both. A second set of points also may have some resonance beyond schools. For we think that the quality constraints on merit rewards discussed here help to explain certain common features in the merit pay plans that we investigated. One such feature was that most merit pay plans were organized around scales with rather small steps. In situations in which everyone thought there was a narrow dispersion of quality among workers, finely graded rewards seemed a way to manage the problem of comparisons. Another common feature was that no district made merit reward depend mainly on classroom performance. Instead, rewards were keyed more heavily to voluntary achievements outside the classroom, in community service and professional accomplishment. In situations in which most teachers thought that they were roughly equally good at teaching, making rewards depend on voluntary work outside of teaching seems a wise

way to manage comparison problems.

Our point is not that these design features are exclusively due to the high and tight quality distribution of teaching. We have mentioned that other circumstances helped to produce these results. And we can easily imagine that the design features just mentioned could occur in districts with mostly poor quality teachers and a tradition of political patronage in schools. But in the districts we studied, it seems likely that these two design features owed something to the particular quality distribution of teaching. It also seems likely that similar quality distributions in other organizations could produce the same result.

The flip side of these points implies caution about certain alternative designs for merit pay schemes. Some practitioners argue that rewards should depend only on classroom quality, but this strikes us as less likely to succeed than more catholic conceptions of merit. Several practitioners also told us that merit pay would work much better if the money incentives were really large, but are skeptical about larger awards. Our analysis suggests that in districts like those studied here, really large awards could be explosive. A complex web of positive sentiments, mutual obligations, and adaptations to circumstance keep good schools running well. It is, in fact, their glue. Trying to make large distinctions of quality in institutions in which the staff correctly sees relatively small differences, could be quite disruptive.

But the most important implication of this part of the story is that merit pay plans in good districts seem to defeat themselves over time. In places where most teachers are good, merit pay will tend to become what one thoughtful administrator in Niskayuna called "an annuity." After considering his district's merit pay scheme, and reviewing merit pay more

generally, he concluded that ". . . both in industry and the public sector, merit assessments tend to become bland and the majority of employees move to the top of the schedules."⁴⁵ The two districts that had the oldest merit pay plans -- Niskayuna and Evanston -- also had the largest fraction of teachers receiving merit awards. But even the district with the newest merit pay plan, Catalina Foothills, had a similarly large fraction of teachers on merit pay.⁴⁶

There is a nice paradox here. Merit pay schemes require distinctions among employees' quality; the idea is to reward the best and inspire the rest. In the districts that we studied, it seems that one reason it was possible to implement such plans was the generally high quality of the teaching staff. A skewed distribution of quality helped to make merit pay politically possible. But over time, this skewed distribution of quality seems to have encouraged a movement toward universal coverage, which tends to erode the foundations of merit pay. To put it a little more directly, the very attributes of design and community that make merit pay successful in the short run tend to make it self-defeating in the long run. This does not mean that some forms of merit pay are not useful under some circumstances. It just means that merit pay is not by itself an enduring or an entirely satisfactory way to strengthen incentives for good teaching.

- 1) Lloyd Young, The Administration of Merit-Type Teachers' Salary Schedules, (New York: Teachers College, 1933), p. 5
- 2) President Reagan said, in May 1983, that "Teachers should be paid and promoted on the basis of their merit and competence. Hard-earned tax dollars should encourage the best."
- 3) See, for example, Edward Lawler, Pay and Organization Development, (Reading, Mass.: Addison-Wesley, 1981), pp. 97-8.
- 4) In addition to Young, E.S. Evenden published a study Teachers' Salaries and Salary Schedules in the United States, 1918-1919, (Washington, D.C: National Education Association, 1919). Elwood P. Cubberly also dealt with the subject in his Portland survey, published as part of Public School Administration (Boston: Houghton Mifflin, 1929). Young mentions a few other studies on pp. 3-5.
- 5) For example, see Young, Merit-type Salary Schedules, pp. 11-23.
- 6) Paul Porwoll, Merit Pay for Teachers, (Arlington, VA: Educational Research Service, 1978), pp. 5, 11-23.
- 7) Frederick Calhoun and Nancy Protheroe, Merit Pay Plans for Teachers, (Arlington, VA: Educational Research Service, 1983), p. 3.
- 8) Porwoll, Merit Pay, p. vi.
- 9) Calhoun and Protheroe, Merit Pay Plans, p. 23.
- 10) VBC.3. We cite all of our interviews by the codes we used to identify them for data analysis purposes.
- 11) VBP. 9
- 12) VBG. 2
- 13) VBS. 1

- 14) VBBM. 11
- 15) VBB. 3
- 16) VBH. 10
- 17) EM. 3
- 18) EM. 13
- 19) EJ. 2
- 20) EM. 1
- 21) EM. 2-3
- 22) EM. 3
- 23) EJ. 2
- 24) EJ. 4
- 25) EA. 1
- 26) EM. 9, 12
- 27) MMP. 11-12
- 28) ML. 1
- 29) MF. 2-3
- 29a) ES. 2
- 30) ML. 4, MB. 2
- 31) VV. 5
- 32) VH. 15
- 33) EB. 14
- 34) MB. 1
- 35) MBP. 2-3
- 36) NT. 1
- 37) NB. 4
- 38) EW. 3
- 39) NG. 4
- 40) NC. 6

- 41) CFH. 2
- 42) CFH. 2
- 43) CFH. 4
- 44) CFF. 2
- 45) NJB. 1
- 46) Not everyone in our districts agrees with this view. The top administrators in Catalina Foothills say that they hope to deal with the problem of grade inflation by creating a new merit pay program. The new program that they plan will somehow make distinctions among good teachers, separating good from very good teachers. It also will put much more money on the differences, perhaps as much as five or more thousand dollars a year. If they pursue this plan, it will be fascinating to follow the story as it unfolds.